



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2007

---

**Introduction. Tome 1, Section 2. De la philologie aux nouveaux médias: éditions de  
textes - linguistique de corpus - analyse informatique du langage**

Glessgen, Martin-Dietrich ; Di Girolamo, Costanzo ; Schøsler, Lene ; Völker, Harald

DOI: <https://doi.org/10.1515/9783110923599.287>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-151256>

Book Section

Published Version

Originally published at:

Glessgen, Martin-Dietrich; Di Girolamo, Costanzo; Schøsler, Lene; Völker, Harald (2007). Introduction. Tome 1, Section 2. De la philologie aux nouveaux médias: éditions de textes - linguistique de corpus - analyse informatique du langage. In: Trotter, David. Actes du XXIVe Congrès International de Linguistique et de Philologie Romanes. Aberystwyth 2004. Tübingen: Niemeyer, 287-290.

DOI: <https://doi.org/10.1515/9783110923599.287>

## Introduction

Le développement des nouveaux médias dans les deux dernières décennies du XX<sup>e</sup> siècle a transformé de façon radicale le «pouvoir explicatif» des sciences textuelles. Cette révolution, à peine entamée, ouvre des champs jusqu'ici inexistantes à la philologie, entendue comme science de l'édition et de l'analyse des textes anciens et modernes (cf. Bozzi/Cignoni/Lebrave 2004). Dernièrement, un lien entre la philologie et la linguistique de corpus vient de se profiler dans les agendas scientifiques (cf. Pusch/Kabatek/Raible 2005). Cherchant à tenir compte de ces possibilités et développements récents, les directeurs de la section 2, intitulée *De la philologie aux nouveaux médias: éditions de textes – linguistique de corpus – analyse informatique du langage*, avaient proposé aux collègues au sein et en dehors de la romanistique d'étudier l'actuel renouvellement méthodologique dans la philologie ainsi que dans les outils informatiques.

Dans une première partie, les contributions étaient centrées sur des *Editions de textes électroniques et traditionnelles*. Les participants étaient censés se prononcer sur les questions suivantes: les éditions – pourquoi? pour qui? suivant quelles méthodes? Les participants de la section ont été invités à réfléchir sur les points suivants:

- 1° l'utilisation des nouvelles méthodologies tant dans l'élaboration de l'édition que dans la présentation de celle-ci (éditions électroniques);
- 2° les témoins concrets, leur médialité et leur élaboration philologique (bases de données, concordanciers etc.);
- 3° les innovations dans la philologie non informatique.

La seconde partie de la section était intitulée *Linguistique de corpus et méthodologies de l'analyse informatique du langage*. Cette partie s'était proposé de contribuer au développement d'une pensée algorithmique qui vise à interpréter le langage à l'aide d'instruments informatiques, tout en tirant profit des acquis de la tradition philologique. Les thématiques traitées ont été les suivantes:

- 1° le balisage et la préparation informatique de corpus textuels en vue d'interrogations variationnelles, historiques et typologiques;
- 2° les interrogations algorithmiques appliquées à des corpus textuels romans, notamment les possibilités de développement d'une syntaxe historique comparée, assistée par l'informatique;
- 3° la critique des outils informatiques disponibles ou en voie de création (finalités linguistiques, utilité, accessibilité).

Les directeurs de la section ont reçu un grand nombre de propositions de contribution. Comme la plupart d'entre elles étaient liées et au premier bloc et au second bloc thématique, il a été décidé d'ignorer la bipartition proposée au début dans la description de la section et de grouper les contributions *grosso modo* dans les trois domaines thématiques suivants, clairement liés entre eux:

- 1° la philologie éditoriale et les éditions;
- 2° la création de bases de données à divers usages;
- 3° les analyses informatiques et l'utilisation des bases de données.

Pour garantir la cohésion thématique de la section, les directeurs de la section ont proposé d'organiser le travail en trois formats:

- 1° les conférences;
- 2° à la fin de chaque journée, un «atelier informatique» suscitant des échanges à la fois plus détaillés et plus informels, basés sur des présentations concernant plus précisément les aspects informatiques;
- 3° comme clôture de la section un atelier méthodologique organisé en table ronde, qui a permis aux participants de revenir sur des questions fondamentales très présentes dans les différentes contributions, et d'en débattre.

Les contributions ont porté sur l'italien, l'espagnol, le portugais, le galicien, l'occitan et le français. Elles ont présenté le fruit de recherches extensives, très souvent menées en équipe. Les communications, ainsi que les débats très animés qu'elles ont soulevés, ont permis de donner un aperçu des grandes interrogations actuelles de la recherche et des orientations les plus porteuses en vue aujourd'hui.

Quelques observations s'imposent. Pour ce qui est de l'utilisation des nouvelles technologies dans notre domaine, il semble utile de distinguer au moins deux phases:

Une première phase au cours de laquelle les chercheurs sont passés du support papier au support électronique. A cette première phase appartient le début de la *Base du Français Médiéval* (BFM, cf. Guillot, Heiden, Lavrentiev), la première compilation du *Corpus d'Amsterdam* (CA), le début du *Dictionnaire du moyen français* à Nancy (DMF) et de l'*Anglo-Norman Dictionary* (AND, cf. Beddow/Rothwell/Trotter).

La seconde phase – dans laquelle nous nous trouvons actuellement – explore les possibilités multiples des nouvelles technologies. Ceci vaut évidemment pour le développement ultérieur des quatre projets mentionnés, pour lesquels se pose tout particulièrement la question de la réutilisation des données et de leur utilisation future. À ceci s'ajoute un autre type de réutilisation, celui de l'intégration de ressources venant d'autres projets, phénomène pertinent pour l'actuel développement du *Dictionnaire du moyen français* au sein du laboratoire ATILF (cf. Souvay) et pour le projet englobant entre autres le *Corpus d'Amsterdam* (cf. Gleßgen). La seconde phase dans l'utilisation des nouvelles technologies a été illustrée de façon impressionnante par la multifonctionnalité de certains projets – non seulement de projets très vastes comme l'*Opera de Vocabolario Italiano* (OVI, cf. Beltrami) et l'état actuel de la BFM, d'ATILF et de l'AND, mais aussi de projets spécialisés comme ceux de Bozzi/Corradini, de Schaller/Uhlig, de Thibault, de Dupuis et al., du *Dicionario de dicionarios do galego medieval* (DDGM; cf. González Seoane), de Peter Ricketts ainsi que le RIALTO (qui a servi de point de départ aux réflexions méthodologiques de Franchi). Il faut également mentionner la visualisation étonnante des importantes analyses dialectométriques de Hans Goebel. Impressionnantes sont également les différentes recherches portant sur les versions et les variantes textuelles (Bozzi/Corradini, Enrique-Arias, Ganascia, Limacher-Riebold et Schaller/Uhlig).

Ces dernières présentations et celles de Carapezza, de Barbosa Machado et de Ricketts, ont alimenté des discussions fondamentales, concernant entre autres les questions suivantes: les bases de données textuelles, «brutes» ou enrichies, et les éditions critiques, sont-elles méthodologiquement compatibles? Et comment combiner ces sources et la lexicographie? Quel est l'apport de l'analyse automatique pour la stemmatologie? Les réponses à ces questions ne sont pas simples, mais la discussion méthodologique avance. Un encodage neutre des données textuelles, de type XML, permet des combinaisons solides

de différentes vues sur les témoins d'une même tradition textuelle. La lexicographie informatique intensifie de plus en plus les liens entre des bases de données textuelles et les bases lexicographiques et interprétatives. Quant à la stemmatologie, les réponses actuelles ne peuvent pas encore satisfaire tous les points de vue. Mais la réflexion est clairement lancée.

Insistons en outre sur le fait novateur que plusieurs projets intègrent de façon élaborée les diasystèmes de la langue (celui de Trèves présenté par Overbeck/Völker et celui de Zurich présenté par Martin-D. Gleßgen et Anne-Christelle Matthey) et, plus généralement, sur l'importance des efforts consacrés à l'élaboration des dictionnaires – même en dehors des vastes projets déjà mentionnés – et enfin sur l'intérêt des aspects typologiques et francophones (André Thibaut) des études de corpus.

Finalement, nous avons abordé la question de l'accessibilité aux sources primaires, par exemple le traitement de l'image (Bozzi/Corradini).

Le but de cette section était de présenter la tradition et les innovations dans le domaine. Nous autres organisateurs tenons à préciser que le point de départ de cette section 2 a été le décalage communicatif entre la «New Philology» et la «vieille» philologie comme cela a été constaté par Peter Dembowski (1998, 401):

C'est pourquoi la *New Philology* a été perçue comme une proclamation du *novus ordo saeculorum*. Dans ce nouvel ordre, les philologues du passé deviennent par nécessité des «dinosaures» nationalistes ou, pire, positivistes. Ce dénigrement du passé blesse ceux qui sans se réclamer complètement de ce passé, tâchent de le comprendre. Le physicien d'aujourd'hui ne dénigre pas Faraday: il le corrige.

Nous sommes heureux de l'intérêt des contributions et des discussions qui ont eu lieu à Aberystwyth. Souhaitant que la collaboration, déjà engagée au sein d'un grand nombre de projets, soit encore plus intense à la suite de cette rencontre, nous espérons avoir contribué à renforcer, comme cela fut déjà – d'une manière encouragente – le cas lors d'un colloque à Jena (cf. Gleßgen/Lebsanft 1997), le dialogue respectueux entre la «vieille» et la «nouvelle» philologie.

## Bibliographie

- Bozzi, Andrea, Laura Cignoni, Jean-Louis Lebrave (ed.) (2004): *Digital Technology and Philological Disciplines*. Pisa/Roma. Istituti Editoriali e Poligrafici Internazionali.
- Dembowski, Peter F. (1998): Les débats américains sur la philologie textuelle de l'ancien français. In: J. Claude Faucon, Alain Labbé, Danielle Quéruel (ed.): *Miscellanea mediaevalia. Mélanges offerts à Philippe Ménard*. Vol. 1: Paris: Champion, 395-405.
- Gleßgen, Martin-D., Franz Lebsanft (ed.) (1997): *Alte und neue Philologie*. Tübingen: Niemeyer.
- Pusch, Claus D., Johannes Kabatek, Wolfgang Raible (ed.) (2005): *Romanistische Korpuslinguistik II: Korpora und diachrone Sprachwissenschaft / Romance Corpus Linguistics II: Corpora and Diachronic Linguistics*. Tübingen: Narr.

